

The Determination of Free ϵ -Amino Groups of Lysine in Proteins Using ^{19}F NMR Spectroscopy¹

J. E. RAMIREZ,² J. R. CAVANAUGH,³
K. S. SCHWEIZER, AND P. D. HOAGLAND

*Eastern Regional Research Center, Agricultural Research Service,
United States Department of Agriculture, Philadelphia, Pennsylvania 19118*

Received April 15, 1974; accepted July 30, 1974

The amount of free ϵ -amino groups of lysine residues in proteins can be determined by a fast and straightforward method based on the use of ^{19}F NMR spectroscopy. Unsubstituted ϵ -amino groups are trifluoroacetylated with the reagent, *S*-ethyl trifluorothioacetate, in dimethyl sulfoxide solution, and the number of such groups is quantitatively determined using standard ^{19}F NMR techniques. The method is evaluated using the following proteins: α -lactalbumin, β -lactoglobulin, bovine serum albumin, β -casein, and lysozyme.

The problem of lysine availability has long been recognized as an important component in the evaluation of the nutritional value of foodstuffs (1). Chemical methods for the determination of available lysine (2–6) are based on the determination of the number of lysine residues with free ϵ -amino groups. When the ϵ -amino group is blocked by derivatization with other substances, the protein chain is inaccessible to hydrolysis by proteolytic enzymes and its nutritive value is lost.

The chemical methods are generally complex and time-consuming. Moreover, recent results (7) with one relatively new method (5) reveal a number of factors affecting the accuracy of the results and leave the general applicability of the method open to some question. We have developed a method for the determination of free ϵ -amino groups of lysine in order to form a basis for the determination of available lysine. It is based on the use of a trifluoroacetylating reagent and the subsequent detection of trifluoroacetylated lysines by ^{19}F NMR spectroscopy. The method as tested with pure proteins offers both reasonable accuracy and a vast improvement in ease of operation.

¹ Presented in part at the 57th Annual Meeting of the Federation of American Societies for Experimental Biology, Atlantic City, NJ, April 15–20, 1973.

² NRC-ARS Postdoctoral Research Associate. Present address: Warner-Lambert, Morris Plains, NJ 07950.

³ Author to whom inquiries should be addressed.

MATERIALS AND METHODS

Dimethyl sulfoxide and *S*-ethyl trifluorothioacetate were obtained from Eastman Kodak Company⁴; bovine serum albumin from Kimex Chemical Company; and lysozyme from Sigma Chemical Company. All were of the highest purity obtainable and were used without further purification. α -Lactalbumin and β -lactoglobulin were gifts from H. Farrell. β -Casein was prepared directly (8).

The ¹⁹F NMR spectra were recorded on a Varian Associates DA-60 IL NMR spectrometer, operating at 56.4 MHz in the high-resolution mode. Relative integrated intensities were determined by the relative weights appropriate to the NMR spectral peaks. The weights were obtained at least in duplicate and were averaged over four or more recordings taken with alternating upfield and downfield sweeps.

Samples were prepared for analysis as follows. One to 1.5 ml of a freshly prepared 2% by weight solution of *S*-ethyl trifluorothioacetate in dimethyl sulfoxide was added to 50–100 mg of protein. The amounts were chosen in a given case to maximize the percentage lysine consistent with the solubility of the trifluoroacetylated protein. Molar ratios of thiol reagent to lysine for all proteins studied fell in the approximate range of 4:1 to 5:1.

When the mixture became homogeneous (on the order of several hours), a portion was added directly to an NMR sample tube and the ¹⁹F NMR spectrum obtained. The percentage of available lysine in the sample is given by:

$$\% \text{ lysine} = \frac{(\text{MW})_{\text{lysine}}}{(\text{MW})_{\text{thiol}}} \times \frac{(^{19}\text{F})_{\text{lysine}}}{(^{19}\text{F})_{\text{total}}} \times \frac{\text{wt solution}}{\text{wt protein}} \times \% \text{ thiol},$$

where (MW)_{lysine} and (MW)_{thiol} refer to the molecular weight of lysine and *S*-ethyl trifluorothioacetate, respectively, (¹⁹F)_{lysine} is the spectral integrated intensity appropriate to the ϵ -amino trifluoroacetyl group, (¹⁹F)_{total} is the total integrated intensity, the weights refer to those of the dimethyl sulfoxide solution and the protein, and % thiol is the % by weight of *S*-ethyl trifluorothioacetate in the dimethyl sulfoxide solution.

RESULTS AND DISCUSSION

The approach used is based on the trifluoroacetylation of the ϵ -amino group of lysine residues and their subsequent quantitation by ¹⁹F NMR spectroscopy. The use of ¹⁹F NMR in conjunction with fluorine probes in proteins was very successfully employed in conformational studies of ribonuclease S (9). There, the method for preparing the trifluoroacetylated lysines in the protein chain was that of Goldberger and Anfinsen (10,11). That technique entailed reacting a 500-fold excess of *S*-ethyl trifluoroacetate with a dilute aqueous solution of protein keeping the pH constant at 10 by the addition of NaOH. The solution was subsequently neutralized, dialyzed, and lyophilized in order to recover the reacted protein. Because the procedure was relatively complex and time-consuming, we sought another method of preparation of the trifluoroacetylated material.

We chose dimethyl sulfoxide as a possible medium for the trifluoroacetylation reaction. Some of the proteins studied are themselves soluble

⁴ The mention of commercial items does not constitute an endorsement by the U. S. Department of Agriculture over others of a similar nature not mentioned.

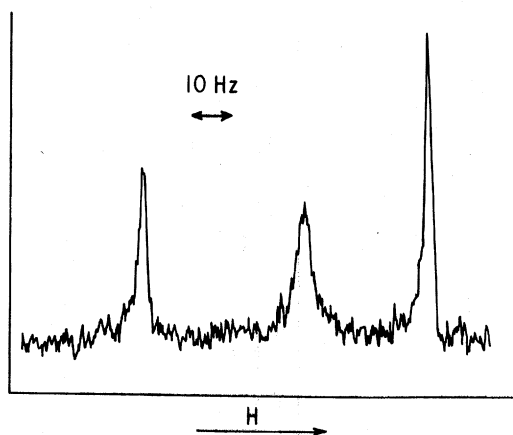


FIG. 1. ^{19}F NMR spectrum of the reaction products of a protein and *S*-ethyl trifluorothioacetate in dimethyl sulfoxide solution. From left to right, the peaks represent trifluoroacetyl groups of trifluoroacetic acid, on the ϵ -amino group of lysine residues, and of unreacted thiol ester.

in dimethyl sulfoxide as is the thiol reagent. Presumably, the reaction would take place much more rapidly in solution than in the immiscible thiol ester–water system where reaction is limited to the two-phase interface. The small volume of reactants, higher concentration of protein, and lower molar ratios of thiol ester to lysine would have significant additional advantages.

The reaction in dimethyl sulfoxide does proceed rapidly. Moreover, even for those proteins with limited solubility, the derivatized product itself was soluble for all proteins studied. When the solution becomes homogeneous (on the order of several hours for the slowest reacting proteins), a portion is added directly to an NMR sample tube and the ^{19}F spectrum is obtained. A typical spectrum is shown in Fig. 1.

The resonance at highest field corresponds to the trifluoroacetyl group of the thiol ester (unreacted reagent); the one at intermediate field to the lysine ϵ -amino substituted one; and the peak at lowest field to trifluoroacetic acid (hydrolysis product of the thiol ester). The ester will continue to react slowly with atmospheric water with the consequent increase in the intensity of the peak at lowest field and decrease in the one at highest field. The relative chemical shifts are concentration dependent; the peak corresponding to trifluoroacetic acid moves to higher field as its concentration increases. In extreme cases in which the ester is fully hydrolyzed, the trifluoroacetic acid peak will lie to the right of the resonance of the trifluoroacetyl group attached to lysine.

The relative integrated intensities were calculated from the weights of the appropriately cut-out peaks determined at least in duplicate. The

molar ratio of lysine to the original thiol ester is given in a straightforward manner by the ratio of the weight of the middle peak to the total weight of all peaks.

In the initial experiments, the quantitative extent of the trifluoroacetylation reaction seemed to be dependent on a number of factors, including the ratio of protein to thiol reagent and the concentration of the reagent. The latter proved to be the most significant factor. In fact, if the thiol concentration were kept constant, the amount of lysine determined was found to be independent of the molar ratio of lysine to thiol. On the other hand, the percentage of lysine determined ran very high (well exceeding 100%) or very low directly as a function of the thiol concentration. Evidently, with a high concentration of thiol reagent, some lysine sites were being double substituted; with a low concentration, not all ϵ -amino groups were being attacked. The optimum concentration for the thiol ester was found to be 2% by weight in dimethyl sulfoxide. Molar ratios of thiol reagent to lysine could vary from 2:1 to 6:1 without affecting the results. A ratio of approximately 4:1 was chosen as suitable.

The α -amino group of the protein chain does not appear to be attacked in these experiments. At least, no resonance in the ^{19}F spectrum could be identified as belonging to a trifluoroacetyl on the α -amino group. The chemical shift difference between trifluoroacetyl groups on α - and ϵ -amino groups is sufficiently large (9) (on the order of 20 Hz at 56.4 MHz) to ensure that if some α -amino groups were being substituted, the corresponding resonance would be distinct and recognizable from that appropriate to the ϵ -amino group.

The spectrum obtained from the reacted lysozyme did show a small

TABLE 1
COMPARISON OF LYSINE CONTENT DETERMINED BY THE NMR METHOD
WITH COMPOSITIONAL DATA

Protein	Lysine content ^a	
	By composition	Experimental
Bovine serum albumin	12.8 ^b	12.5 \pm .4
β -Casein	6.5 ^c	5.9 \pm .1
α -Lactalbumin	11.4 ^c	11.7 \pm .5
β -Lactoglobulin	11.4 ^c	11.5 \pm .4
Lysozyme	5.7 ^b	5.0 \pm .1

^a Grams of lysine per 100 g protein.

^b Reference 12.

^c Reference 13.

spurious upfield resonance that could possibly be due to the trifluoroacetylated ϵ -amino group of doubly trifluoroacetylated N terminal lysine residues [cf. results on ribonuclease S-peptide (9)]. The absence of a peak corresponding to the trifluoroacetylated α -amino group casts some doubt on this interpretation, but the signal-to-noise ratio is sufficiently poor in these spectra that a small resonance could be lost in the noise. This question could probably be answered with the help of improved sensitivity.

The results obtained with a number of purified proteins are shown in Table 1. The error limits on the experimental values are average deviations and represent simply the reproducibility of at least four complete analyses made with each protein. The agreement between compositional data and experimental is excellent. The values for β -casein and lysozyme are somewhat low but are well within the accuracy demanded of a determination of this kind. We attempted to use another method for determining available lysine (5) for lysozyme as a comparison, but we were unable to achieve a precision approaching the NMR method.

The present method, therefore, offers a reasonably accurate, fast, and relatively simple procedure, without the difficulty of variability of results from varying amounts of the protein present. It does require access to a ^{19}F NMR spectrometer and fairly concentrated protein solutions. The latter problem can be largely eliminated by the use of Fourier-Transform pulsed NMR which is rapidly becoming more widely available.

REFERENCES

1. MAURON, J. (1961) in Conference on Meeting Protein Needs of Infants and Preschool Children, Publication 843, p. 425, Nat. Acad. Sci., Nat. Res. Council, Washington, D. C.
2. CARPENTER, K. J. (1960) *Biochem. J.* **77**, 604.
3. BLOM, L., HENDRICKS, P., AND CARIS, J. (1967) *Anal. Biochem.* **21**, 382.
4. ROACH, A. G., SANDERSON, P., AND WILLIAMS, D. R. (1967) *J. Sci. Fd. Agric.* **18**, 274.
5. KAKADE, M. L., AND LIENER, I. E. (1969) *Anal. Biochem.* **27**, 273.
6. FINLEY, J. W., AND FRIEDMAN, M. (1973) *Cereal Chem.* **50**, 101.
7. POSATI, L. P., HOLSINGER, V. H., DE VILBISS, E. D., AND PALLANSCH, M. J. (1972) *J. Dairy Sci.* **55**, 1660.
8. ASCHAFFENBURG, R. (1963) *J. Dairy Res.* **30**, 259.
9. HUESTIS, W. H., AND RAFTERY, M. A. (1971) *Biochemistry* **10**, 1181.
10. GOLDBERGER, R. F., AND ANFENSEN, C. B. (1962) *Biochemistry* **1**, 401.
11. GOLDBERGER, R. F. (1967) *Methods Enzymol.* **11**, 317.
12. TRISTRAM, G. R. (1953) in *The Proteins* (H. Neurath and K. Bailey, eds.), Vol. 1, Part A, p. 181, Academic Press, New York.
13. JENNESS, R., AND PATTON, S. (1959) *Principles of Dairy Chemistry*, p. 125, John Wiley & Sons, New York.